SEQUENCE LISTING

<110> Tarr, Phillip I
     Bilge, Sima S
     Besser, Thomas E
     Vary Jr., James C

<120> ESCHERICHIA COLI O157:H7 EPITHELIAL ADHESIN

<130> WRFO-1-15124

<140>
<141>

<150> 09/098,082
<151> 1998-06-16

<150> 08/765,081
<151> 1997-03-26

<150> PCT/US95/06994
<151> 1995-06-07

<150> 08/265,714
<151> 1994-06-24

<160> 12

<170> PatentIn Ver. 2.1

<210> 1
<211> 8041
<212> DNA
<213> Escherichia coli

<220>
<221> misc feature
<222> (1) . . . (8041)
<223> nucleic acid molecule wherein n represents a, c, g or t

<400> 1
```
ggtacctgtc gccagttctc cgatctgttt acccgggaaa ttatctccta cagcttgtca 60
gaaaggtcgg tgatggagca tcgntaatac gatgctatac gatgnattca cagtgcccgg 120
ccagaggatg ccccgtcgct gcatatggat cagngttggc aatatcgaat tgcaggctat 180
aggcaaagtt aangccccat ggagtagcac aaaatatgcc gcacanagga aacggtctga 240
ataacgcagt gatgaagaac ttcttcagca cactgctaaa acgcatagtg atcgagcgct 300
gattctggcg aacaactgaa ctaatacatc agaatctgca ttatgttaaa taaatataaa 360
aagatggttt aaataccccg ttacttgtga cttacactat acggtatcgc atcgtttaat 420
attcgcaccg gccagatttt tatttctatt agttgtcaca atactgaatg cgtacgacca 480
cagtattctg gctcctgtgt ggttatgctt taattctgcg ttccgggcag ataagcagtt 540
gcttgcagga atccttcttg tgttaatgtc agttcccctt ttaccagtgc tgatttccac 600
attccgtcca acagagctta tagcctttcc ctggattata gcattgtccg gctgaagttc 660
tttttgaata ataatagaag cactgctggc agatccagtc cgttttttcat aacccactgt 720
actgataacc ataatctaat cagtagaaat tgagtcgaaa ataagcacta ctccatacag 780
gataattaga ggtcagtttg attattcaca attcatcatc agcattttct atttctgacg 840
aaatcaatat gaaaataacc atatatgata attattataa taacggcttt aattggaata 900
catatattac aacgtattat atataattgg tattctggga actatattct caaaatacag 960
tagaaacgag gtatgtttct ggtggaaagg acagtgggat taaaaagtaa gaattgataa 1020
aaaaacgcca gcaacacata gtgccggngg agggaatacc ccatggagaa aatgtgatgc 1080
```

-1-

```
ctagagcatc atgantatat aattaaaaat agttagcgtt gtcactactt cnacaaaaat 1140
aattttcgta gtatagaaag atatttttat gcatgaccta cctgaatttg ctccgggtag 1200
aggttataaa taaaaattga atcacgacaa acacaatatt catagtatgg cgatgcctac 1260
gccagcaaga atagcgncaa taatattggg aatataatag acaccagacg cacaggcatc 1320
tcactcctta acaaacaaca atcaggatat ttacttttac caagctaact gtttacaccc 1380
aaagtacaca caattaacca ttcaataaca aatgncaata tccatagcca tacgacttta 1440
ccttgtaatg ttcggtattt ctttataatt attctgggaa atctaacatt tattttttaaa 1500
atcaaattat cttgttgttt aaaaataagt tcacatactt tatcatcttc ttccgccatc 1560
aacatctctg catacttaaa catttcagaa cgttccttta gcacagaaga gtaattatat 1620
gtccagttcc caggcaataa tgcttatgga tatttaattc ataatttaga gaatattttg 1680
caataatatt tggcagtatt cagaaatacc tgaaaaatca tactatacag ccctagggaa 1740
tggataaaga ttctaacaaa gcatttcaac aatatatact tgttaaaaat ccatatcgaa 1800
atgccgttgc agcattaata tatgcctcat tcataaaatg taaaagagca agctcgtacc 1860
agtaggggggg aaattcattg acatgtcctg tcaatacgta cagagccctg ccatgcttgc 1920
cagcacgtaa caattccgcc cccagtaacc agcgtgttcc ctgattattc tcagggttat 1980
atgttagtat tttatcaatg agtataacgg catcctgatg acgctgtaaa tgaacattcg 2040
ccagaattgc agcataaaga gcgcaaatgc atatgccaga tgagcgtgaa tatcaatgat 2100
gtcaggagca tgtcaggcaa gacgcctgag tttgttgacg taatttgttt cgtttatttc 2160
atcactatcg tatgcgtcaa gagttttttc aaactcatcc cgaaatgggc caagcctggt 2220
ttcaggaaat aagaaatacc cttggttatt gttcacttaa aagtataatt tgtaattcaa 2280
atcctgagcc tgtaaagcgg ggagtaacat attctattct gaagagaata aaagtcgtga 2340
tgcgatgtat caagcccgga ttgtaatccc agattaacat agatcacaaa acaacttatt 2400
cttcactaac gtcaagataa atatcgttgt atgccttatc acgactacga ataccagcaa 2460
gatacaactg acatacggaa agataccacg tttttttactc ccgaaaataa cgctaaaaag 2520
ctacttcccc atcgtttgtc cttagtattg ccagcgccaa caatgtgggc tgacatgata 2580
aagctgtcta ggaaattgtt cgcctcctca gcggacaatc caaatggtga ttgtctctgt 2640
taaacgttta ttttgaaggt cgactgaata aggtgatgac gctgtagaat ttttcacgtg 2700
ccacagaatt ttgaacgctt tctcttacaa tatttcaatg tttctatcag tattcgccgg 2760
aagaagtcat cgaccaaatc atcccagtcg tctcgcatca ctgaccattc atggttgaca 2820
tgtggtggaa atccgcttct acagtaacca tttttttattc gcaaaaccga acacgccata 2880
cggatagctg ttaactggca tgccgaggca gtcgttattt atatttggtt ttgtcaataa 2940
tctttatttt ttgtaaaagg caaatataaa ttattctcat tattgtttgt atttgtgtat 3000
tgtcttgccg gttaacatga tcggagatta gtaatatgcg aataaccact ctggcttccg 3060
tagtcattcc ctgtctcgga ttttcagcca gcagcatagc tgctgcagag gatgtgatga 3120
ttgtctcggc atccggctat gagaaaaagc tgactaacgc agccgccagt gtttctgtga 3180
ttagccagga ggaattgcag tccagccagt accacgatct ggcggaggct ctgagatcag 3240
tagagggtgt ggatgttgaa agtggtacgg gtaaaaccgg agggctggaa atcagcatcc 3300
gaggaatgcc agccagttac acgctgatac tgattgatgg tgttcgtcag ggcggaagca 3360
gtgacgtgac tcccaacggt ttttctgcca tgaataccgg gttcatgccc cctctggccg 3420
ccattgagcg tattgaggtt atcaggggggc cgatgtccac actgtatggc tctgatgcga 3480
tgggcggtgt ggtgaatatc attaccagaa agaatgcaga caaatggctc tcttccgtca 3540
atgcagggct gaatctgcag gaaagcaaca aatgggggtaa cagcagccag tttaatttct 3600
ggagcagtgg tccccttgtg gatgattctg tcagcctgca ggtacgcggt agcacacaac 3660
agcgtcaggg ttcatcggtc acatcactga gcgatacagc aggcacgcgt attccttatc 3720
ccacggagtc acagaattat aatcttggtg cacgtcttga ctggaaggcg tcggagcagg 3780
atgtgctctg gtttgatatg gataccaccc ggcagcgtta tgataaccgg gatgggcaac 3840
tggggagtct gacggggggga tatgaccgga ccctgcgcta tgagcgaaac aaaatttcag 3900
ctggctatga tcatactttc accttcggaa catgaaatc gtatctgaac tggaacgaga 3960
cagaaaataa aggtcgtgag cttgtacgca gtgtactgaa gcgcgacaaa tggggggcttg 4020
ccggtcagcc gcgggagctt aaggaatcga accttatcct gaattcatta ctgcttaccc 4080
ctctgggaga atctcatctg gttacggtgg ggggcgagtt tcagagctcg tccatgaaag 4140
acggagttgt ccttgccagc acaggtgaaa ctttccggca gaaaagctgg tcggtatttg 4200
ctgaggatga gtggcatctc acggatgcac ttgcgctgac tgcgggcagc cgctatgaac 4260
atcatgagca attcgggggga cacttcagtc cgcgtgcata tctggtctgg gatgtggcag 4320
atgcctggac gctgaaaggc ggtgtgacca cgggatataa ggcacccaga atggggcagc 4380
tacataaagg gattagtggt gtgtccgggc agggaaaaac aaatctactt ggtaaccccg 4440
acctgaagcc ggaagagagc gtcagttatg aggctggggt gtattacgat aaccccgccg 4500
gtctgaatgc caatgtcaca ggttttatga ctgacttctc caacaagatt gtctcttatt 4560
ccataaatga taacaccaat agctatgtaa acagcggaaa ggcccggttg cacggtgtgg 4620
```

```
aatttgccgg cacattgccg ctgtggtcag aggatgtcac gctgtcactg aattacacct 4680
ggacccgaag tgaacaacgt gatggtgata acaaaggtgc gccgctgagt tatacccctg 4740
aacacatggt gaatgcgaaa ctgaactggc agatcaccga agaggtggca tcatggctgg 4800
gtgcccgtta tcgcgggaaa acaccacgtt tcacccagaa ttattcgtca ctgagcgctg 4860
tacagaagaa agtgtatgat gagaaaggag aatacctgaa agcctggacg gtggtggatg 4920
caggtctgtc gtggaagatg acggatgccc tgacgctgaa tgctgcggtg aataacctgc 4980
tcaacaagga ttacagtgac gtgagcctgt acagtgccgg taagagtacg ctgtatgccg 5040
gtgattactt ccagacggga tcatcaacaa caggatatgt gatacctgag cgaaattact 5100
ggatgtcgct gaactatcag ttctgataat aacaaaacgc tatcactgac ggtagaatac 5160
gttgccactg caactcctgg cggaacagtg gcaacgtntt aggttaagtg catttccgat 5220
ccgctaatga gatttcgtta ccaacaacta atatcgtcac aggaaatgca cggattattt 5280
ttaacttatc atttacatac ttgtccagag tgtnagcgca ccgcgacgga cgtggggtaa 5340
aaattagttt acagagagag tgacgttcca ggggaacaac tctttcatgc ggttggcagg 5400
ccaggtgttg gttacactga tcacgtgggc gttggccacg tttccggntc gattccgtta 5460
agttttggag ctaccgatca ggctgtacat cactgncgca ctatcgctcg tcatctcaaa 5520
gtcctgtctc gtcagcagga aggtatcatt ctctcccgcc attttccag gggnccggtc 5580
agataagtcc ctttgtctat cgctgactcc tgactcataa cccggttagc agaatgcagg 5640
ntcaccactc gccacgacca aatccaaata agtcaattgc accttctcaa tcgccatttt 5700
gtcagtaagc gtacagcctc aactgatggt atcttcacca tcaatgacaa cggtgatcgc 5760
aattttactg acgttcgccg gaacacgatc cagtgctatc tcaatgctgg cctgctgcga 5820
accggtaacg agcctgactg cccctcagg agaagacaaa ttattataaa agataaagtc 5880
agaatcgcca ctgacctttc cctgagcatt aagcatgaac agggaggtat cgggttcgcc 5940
ttaaaagccg gatttgncag ggtactgaag attcagcctg atcgnagatt gctgaagggg 6000
tatgtattgt ccggattgta aattcatatt aactctcctg atttntgatt attattaatg 6060
cgcagcgttt atatatgttc cctaggctta gttctggacg ctggatattc ggtgaggcgt 6120
aaatatggta tgacaccatt tttcataacg ctgaagtttc tatacctgtt gaatttgaat 6180
tttcattgac cgggtatctt attttccagg gcccgttcct tcataagtcg caaaagtaac 6240
atatatccga agggcatgct gttgatatca gacacggaat actggcttta accagcaacc 6300
acagataaac ccggggcctg cataaaaggt tcatgccaga attaacatag ctctgctttc 6360
tgcccccccc tttcccatat tcacccgttg atagcggatc gatacccaaa aaaaaccccg 6420
gcattgccgg ggtcagactc agcgtctttt agacattgac gccgtgctgg gaagcaagcg 6480
ctgacagacc accggcaaaa cnctggccaa cagctttaaa cttccactca gtaccatggc 6540
gatacagttc accgaagacc attgcggttt cggttgaggc gtcttcagac agatcgaaac 6600
gggcaatttc cgtcccgttg tcgttgttgt aaacgcgcat gaagctgttg ctcaccatgc 6660
cgaagttttg tttacgcgct tctgcatcat agatggtaac ggcaaatacc agttttttga 6720
tgtctgctga cctttggtc agatcgattt tgacctgctc atcgtcgccg tcgccttcac 6780
cggtacggtt gtcgccctgg tgctctactg cgccatcagg gctggtttta ttattgaaga 6840
aaatgaaatg ggcatctgac agtactttac cgtcttcacc tactgcgaat acggaagcgt 6900
ccagancaaa accctgacca tcggttacac gggcatccca gcccaggcca accatagcga 6960
cattcatggt tggtgcttct ttggtcagag atacgttgcc gccttttacg agagaaactg 7020
ccatttttag ctcctgcaaa cagntgaatg aggctgaata acaccccag aaatgaaaag 7080
ttacttttcg atcaggacgc gttaatnccg tactgagcac atacagatgc cagaccacca 7140
gcataacnct gtnctactgc gcggaatttc cactcaccat tgtggcgaga cagctcgccg 7200
cgcagcatgg cagtctcagt ggacgcatct tcggtcagat cgtagcgagc gacttcagtc 7260
tggttatcgt cattaaccag acgaataaac gcaccggata cctgaccaca gctctggcga 7320
cgagcctgag catcgtggat ggtcacaacg aagatgatct tgtcaacttc agacgggacg 7380
gcgtccagtt taattttcag cgattcatca tcaccatcgc cctcaccggt gcggttatcg 7440
ccggtgtgcg ttacggaacc gtcggatgac gtcaggttgt tatagaagat gaaatctgaa 7500
tcgccgcgca ctttgccgtt tgaggccagc aggaatgctg aagcatccag gtcaaagtcc 7560
tgaccgtctg ttgaacgcgc atcccagcca aggcccacca ggacattttt cattgacgga 7620
gctgctttac tcagggagac gttcccgcct gtggaaagag aaacactcat aaaataccct 7680
cttcgattag taattgttca ggttaacact taaggggatt atctcccctt ttcctcagat 7740
tcaggtgtgc ccgggaacat gacgcttgcg agaatgccca gcgccagtac acccagtaca 7800
acatacaggc tggttgttgc cgcgatgctg taaccatgat gccagatgtg atcgatcgca 7860
ttcaggccga gttttgccac gatgaagaac agcancacga tagcggcctt ctccagatgn 7920
accaggtnct gtttcagtgc ctcnaggaca aaatacagag tacgcagacc caggatagca 7980
aacatcatgg cactatagac gatgaagcgg ttcacgactg acggcaatga tttccggtac 8040
c                                                                8041
```

<210> 2
<211> 574
<212> DNA
<213> Escherichia coli

<400> 2
atggcagttt ctctcgtaaa aggcggcaac gtatctctga ccaaagaagc accaaccatg 60
aatgtcgcta tggttggcct gggctgggat gcccgtgtaa ccgatggtca gggttttgtc 120
tggacgcttc cgtattcgca gtaggtgaag acggtaaagt actgtcagat gcccatttca 180
ttttcttcaa taataaaacc agccctgatg gcgcagtaga gcaccagggc gacaaccgta 240
ccggtgaagg cgacggcgac gatgagcagg tcaaaatcga tctgaccaaa gtctcagcag 300
acatcaaaaa actggtattt gccgttacca tctatgatgc agaagcgcgt aaacaaaact 360
tcggcatggt gagcaacagc ttcatgcgcg tttacaacaa cgacaacggg acggaaattg 420
cccgtttcga tctgtctgaa gacgcctcaa ccgaaaccgc aatggtcttc ggtgaactgt 480
atcgccatgg tactgagtgg aagtttaaag ctgttggcca ggtttttgccg gtggtctgtc 540
agcgcttgct tcccagcacg gcgtcaatgt ctaa 574

<210> 3
<211> 576
<212> DNA
<213> Escherichia coli

<400> 3
atgagtgttt ctctttccac aggcgggaac gtctccctga gtaaagcagc tccgtcaatg 60
aaaaatgtcc tggtggggcct tggctgggat gcgcgttcaa cagacggtca ggactttgac 120
ctggatgctt cagcattcct gctggcctca aacggcaaag tgcgcggcga ttcagatttc 180
atcttctata caacctgac gtcatccgac ggttccgtaa cgcacaccgg cgataaccgc 240
accggtgagg gcgatggtga tgatgaatcg ctgaaaatta aactggacgc cgtcccgtct 300
gaagttgaca agatcatctt cgttgtgacc atccacgatg ctcaggctcg tcgccagagc 360
tgtggtcagg tatccggtgc gtttattcgt ctggttaatg acgataacca gactgaagtc 420
gctcgctacg atctgaccga agatgcgtcc actgagactg ccatgctgcg cggcgagctg 480
tctcgccaca atggtgagtg gaaattccgc gcagtagaca ggttatgctg gtggtctggc 540
atctgtatgt gctcagtacg gattaacgcg tcctga 576

<210> 4
<211> 2091
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(2088)

<400> 4
atg cga ata acc act ctg gct tcc gta gtc att ccc tgt ctc gga ttt 48
Met Arg Ile Thr Thr Leu Ala Ser Val Val Ile Pro Cys Leu Gly Phe
1   5   10   15

tca gcc agc agc ata gct gct gca gag gat gtg atg att gtc tcg gca 96
Ser Ala Ser Ser Ile Ala Ala Ala Glu Asp Val Met Ile Val Ser Ala
            20   25   30

tcc ggc tat gag aaa aag ctg act aac gca gcc gcc agt gtt tct gtg 144
Ser Gly Tyr Glu Lys Lys Leu Thr Asn Ala Ala Ala Ser Val Ser Val
            35   40   45

```
att agc cag gag gaa ttg cag tcc agc cag tac cac gat ctg gcg gag     192
Ile Ser Gln Glu Glu Leu Gln Ser Ser Gln Tyr His Asp Leu Ala Glu
        50                  55                  60

gct ctg aga tca gta gag ggt gtg gat gtt gaa agt ggt acg ggt aaa     240
Ala Leu Arg Ser Val Glu Gly Val Asp Val Glu Ser Gly Thr Gly Lys
    65                  70                  75                  80

acc gga ggg ctg gaa atc agc atc cga gga atg cca gcc agt tac acg     288
Thr Gly Gly Leu Glu Ile Ser Ile Arg Gly Met Pro Ala Ser Tyr Thr
                    85                  90                  95

ctg ata ctg att gat ggt gtt cgt cag ggc gga agc agt gac gtg act     336
Leu Ile Leu Ile Asp Gly Val Arg Gln Gly Gly Ser Ser Asp Val Thr
                100                 105                 110

ccc aac ggt ttt tct gcc atg aat acc ggg ttc atg ccc cct ctg gcc     384
Pro Asn Gly Phe Ser Ala Met Asn Thr Gly Phe Met Pro Pro Leu Ala
            115                 120                 125

gcc att gag cgt att gag gtt atc agg ggg ccg atg tcc aca ctg tat     432
Ala Ile Glu Arg Ile Glu Val Ile Arg Gly Pro Met Ser Thr Leu Tyr
        130                 135                 140

ggc tct gat gcg atg ggc ggt gtg gtg aat atc att acc aga aag aat     480
Gly Ser Asp Ala Met Gly Gly Val Val Asn Ile Ile Thr Arg Lys Asn
145                 150                 155                 160

gca gac aaa tgg ctc tct tcc gtc aat gca ggg ctg aat ctg cag gaa     528
Ala Asp Lys Trp Leu Ser Ser Val Asn Ala Gly Leu Asn Leu Gln Glu
                165                 170                 175

agc aac aaa tgg ggt aac agc agc cag ttt aat ttc tgg agc agt ggt     576
Ser Asn Lys Trp Gly Asn Ser Ser Gln Phe Asn Phe Trp Ser Ser Gly
                180                 185                 190

ccc ctt gtg gat gat tct gtc agc ctg cag gta cgc ggt agc aca caa     624
Pro Leu Val Asp Asp Ser Val Ser Leu Gln Val Arg Gly Ser Thr Gln
            195                 200                 205

cag cgt cag ggt tca tcg gtc aca tca ctg agc gat aca gca ggc acg     672
Gln Arg Gln Gly Ser Ser Val Thr Ser Leu Ser Asp Thr Ala Gly Thr
    210                 215                 220

cgt att cct tat ccc acg gag tca cag aat tat aat ctt ggt gca cgt     720
Arg Ile Pro Tyr Pro Thr Glu Ser Gln Asn Tyr Asn Leu Gly Ala Arg
225                 230                 235                 240

ctt gac tgg aag gcg tcg gag cag gat gtg ctc tgg ttt gat atg gat     768
Leu Asp Trp Lys Ala Ser Glu Gln Asp Val Leu Trp Phe Asp Met Asp
                245                 250                 255

acc acc cgg cag cgt tat gat aac cgg gat ggg caa ctg ggg agt ctg     816
Thr Thr Arg Gln Arg Tyr Asp Asn Arg Asp Gly Gln Leu Gly Ser Leu
                260                 265                 270

acg ggg gga tat gac cgg acc ctg cgc tat gag cga aac aaa att tca     864
Thr Gly Gly Tyr Asp Arg Thr Leu Arg Tyr Glu Arg Asn Lys Ile Ser
        275                 280                 285
```

```
gct ggc tat gat cat act ttc acc ttc gga aca tgg aaa tcg tat ctg      912
Ala Gly Tyr Asp His Thr Phe Thr Phe Gly Thr Trp Lys Ser Tyr Leu
    290             295             300

aac tgg aac gag aca gaa aat aaa ggt cgt gag ctt gta cgc agt gta      960
Asn Trp Asn Glu Thr Glu Asn Lys Gly Arg Glu Leu Val Arg Ser Val
305             310             315             320

ctg aag cgc gac aaa tgg ggg ctt gcc ggt cag ccg cgg gag ctt aag     1008
Leu Lys Arg Asp Lys Trp Gly Leu Ala Gly Gln Pro Arg Glu Leu Lys
            325             330             335

gaa tcg aac ctt atc ctg aat tca tta ctg ctt acc cct ctg gga gaa     1056
Glu Ser Asn Leu Ile Leu Asn Ser Leu Leu Leu Thr Pro Leu Gly Glu
            340             345             350

tct cat ctg gtt acg gtg ggg ggc gag ttt cag agc tcg tcc atg aaa     1104
Ser His Leu Val Thr Val Gly Gly Glu Phe Gln Ser Ser Ser Met Lys
        355             360             365

gac gga gtt gtc ctt gcc agc aca ggt gaa act ttc cgg cag aaa agc     1152
Asp Gly Val Val Leu Ala Ser Thr Gly Glu Thr Phe Arg Gln Lys Ser
        370             375             380

tgg tcg gta ttt gct gag gat gag tgg cat ctc acg gat gca ctt gcg     1200
Trp Ser Val Phe Ala Glu Asp Glu Trp His Leu Thr Asp Ala Leu Ala
385             390             395             400

ctg act gcg ggc agc cgc tat gaa cat cat gag caa ttc ggg gga cac     1248
Leu Thr Ala Gly Ser Arg Tyr Glu His His Glu Gln Phe Gly Gly His
            405             410             415

ttc agt ccg cgt gca tat ctg gtc tgg gat gtg gca gat gcc tgg acg     1296
Phe Ser Pro Arg Ala Tyr Leu Val Trp Asp Val Ala Asp Ala Trp Thr
            420             425             430

ctg aaa ggc ggt gtg acc acg gga tat aag gca ccc aga atg ggg cag     1344
Leu Lys Gly Gly Val Thr Thr Gly Tyr Lys Ala Pro Arg Met Gly Gln
            435             440             445

cta cat aaa ggg att agt ggt gtg tcc ggg cag gga aaa aca aat cta     1392
Leu His Lys Gly Ile Ser Gly Val Ser Gly Gln Gly Lys Thr Asn Leu
    450             455             460

ctt ggt aac ccc gac ctg aag ccg gaa gag agc gtc agt tat gag gct     1440
Leu Gly Asn Pro Asp Leu Lys Pro Glu Glu Ser Val Ser Tyr Glu Ala
465             470             475             480

ggg gtg tat tac gat aac ccc gcc ggt ctg aat gcc aat gtc aca ggt     1488
Gly Val Tyr Tyr Asp Asn Pro Ala Gly Leu Asn Ala Asn Val Thr Gly
            485             490             495

ttt atg act gac ttc tcc aac aag att gtc tct tat tcc ata aat gat     1536
Phe Met Thr Asp Phe Ser Asn Lys Ile Val Ser Tyr Ser Ile Asn Asp
            500             505             510

aac acc aat agc tat gta aac agc gga aag gcc cgg ttg cac ggt gtg     1584
Asn Thr Asn Ser Tyr Val Asn Ser Gly Lys Ala Arg Leu His Gly Val
```

```
gaa ttt gcc ggc aca ttg ccg ctg tgg tca gag gat gtc acg ctg tca       1632
Glu Phe Ala Gly Thr Leu Pro Leu Trp Ser Glu Asp Val Thr Leu Ser
        530                     535                     540

ctg aat tac acc tgg acc cga agt gaa caa cgt gat ggt gat aac aaa       1680
Leu Asn Tyr Thr Trp Thr Arg Ser Glu Gln Arg Asp Gly Asp Asn Lys
545                     550                     555                 560

ggt gcg ccg ctg agt tat acc cct gaa cac atg gtg aat gcg aaa ctg       1728
Gly Ala Pro Leu Ser Tyr Thr Pro Glu His Met Val Asn Ala Lys Leu
                565                     570                     575

aac tgg cag atc acc gaa gag gtg gca tca tgg ctg ggt gcc cgt tat       1776
Asn Trp Gln Ile Thr Glu Glu Val Ala Ser Trp Leu Gly Ala Arg Tyr
                580                     585                     590

cgc ggg aaa aca cca cgt ttc acc cag aat tat tcg tca ctg agc gct       1824
Arg Gly Lys Thr Pro Arg Phe Thr Gln Asn Tyr Ser Ser Leu Ser Ala
            595                     600                     605

gta cag aag aaa gtg tat gat gag aaa gga gaa tac ctg aaa gcc tgg       1872
Val Gln Lys Lys Val Tyr Asp Glu Lys Gly Glu Tyr Leu Lys Ala Trp
        610                     615                     620

acg gtg gtg gat gca ggt ctg tcg tgg aag atg acg gat gcc ctg acg       1920
Thr Val Val Asp Ala Gly Leu Ser Trp Lys Met Thr Asp Ala Leu Thr
625                     630                     635                 640

ctg aat gct gcg gtg aat aac ctg ctc aac aag gat tac agt gac gtg       1968
Leu Asn Ala Ala Val Asn Asn Leu Leu Asn Lys Asp Tyr Ser Asp Val
                645                     650                     655

agc ctg tac agt gcc ggt aag agt acg ctg tat gcc ggt gat tac ttc       2016
Ser Leu Tyr Ser Ala Gly Lys Ser Thr Leu Tyr Ala Gly Asp Tyr Phe
                660                     665                     670

cag acg gga tca tca aca aca gga tat gtg ata cct gag cga aat tac       2064
Gln Thr Gly Ser Ser Thr Thr Gly Tyr Val Ile Pro Glu Arg Asn Tyr
                675                     680                     685

tgg atg tcg ctg aac tat cag ttc tga                                   2091
Trp Met Ser Leu Asn Tyr Gln Phe
        690                     695
```


<210> 5
<211> 696
<212> PRT
<213> Escherichia coli

<400> 5
```
Met Arg Ile Thr Thr Leu Ala Ser Val Val Ile Pro Cys Leu Gly Phe
 1               5                   10                  15

Ser Ala Ser Ser Ile Ala Ala Ala Glu Asp Val Met Ile Val Ser Ala
                20                  25                  30
```

```
Ser Gly Tyr Glu Lys Lys Leu Thr Asn Ala Ala Ala Ser Val Ser Val
        35                  40                  45

Ile Ser Gln Glu Glu Leu Gln Ser Ser Gln Tyr His Asp Leu Ala Glu
        50                  55                  60

Ala Leu Arg Ser Val Glu Gly Val Asp Val Glu Ser Gly Thr Gly Lys
65                  70                  75                  80

Thr Gly Gly Leu Glu Ile Ser Ile Arg Gly Met Pro Ala Ser Tyr Thr
                85                  90                  95

Leu Ile Leu Ile Asp Gly Val Arg Gln Gly Gly Ser Ser Asp Val Thr
            100                 105                 110

Pro Asn Gly Phe Ser Ala Met Asn Thr Gly Phe Met Pro Pro Leu Ala
        115                 120                 125

Ala Ile Glu Arg Ile Glu Val Ile Arg Gly Pro Met Ser Thr Leu Tyr
    130                 135                 140

Gly Ser Asp Ala Met Gly Gly Val Val Asn Ile Ile Thr Arg Lys Asn
145                 150                 155                 160

Ala Asp Lys Trp Leu Ser Ser Val Asn Ala Gly Leu Asn Leu Gln Glu
            165                 170                 175

Ser Asn Lys Trp Gly Asn Ser Ser Gln Phe Asn Phe Trp Ser Ser Gly
            180                 185                 190

Pro Leu Val Asp Asp Ser Val Ser Leu Gln Val Arg Gly Ser Thr Gln
        195                 200                 205

Gln Arg Gln Gly Ser Ser Val Thr Ser Leu Ser Asp Thr Ala Gly Thr
        210                 215                 220

Arg Ile Pro Tyr Pro Thr Glu Ser Gln Asn Tyr Asn Leu Gly Ala Arg
225                 230                 235                 240

Leu Asp Trp Lys Ala Ser Glu Gln Asp Val Leu Trp Phe Asp Met Asp
            245                 250                 255

Thr Thr Arg Gln Arg Tyr Asp Asn Arg Asp Gly Gln Leu Gly Ser Leu
        260                 265                 270

Thr Gly Gly Tyr Asp Arg Thr Leu Arg Tyr Glu Arg Asn Lys Ile Ser
        275                 280                 285

Ala Gly Tyr Asp His Thr Phe Thr Phe Gly Thr Trp Lys Ser Tyr Leu
    290                 295                 300

Asn Trp Asn Glu Thr Glu Asn Lys Gly Arg Glu Leu Val Arg Ser Val
305                 310                 315                 320

Leu Lys Arg Asp Lys Trp Gly Leu Ala Gly Gln Pro Arg Glu Leu Lys
            325                 330                 335

Glu Ser Asn Leu Ile Leu Asn Ser Leu Leu Leu Thr Pro Leu Gly Glu
            340                 345                 350
```

Ser His Leu Val Thr Val Gly Gly Glu Phe Gln Ser Ser Ser Met Lys
        355   360   365

Asp Gly Val Val Leu Ala Ser Thr Gly Glu Thr Phe Arg Gln Lys Ser
        370   375   380

Trp Ser Val Phe Ala Glu Asp Glu Trp His Leu Thr Asp Ala Leu Ala
385   390   395   400

Leu Thr Ala Gly Ser Arg Tyr Glu His His Glu Gln Phe Gly Gly His
            405   410   415

Phe Ser Pro Arg Ala Tyr Leu Val Trp Asp Val Ala Asp Ala Trp Thr
            420   425   430

Leu Lys Gly Gly Val Thr Thr Gly Tyr Lys Ala Pro Arg Met Gly Gln
        435   440   445

Leu His Lys Gly Ile Ser Gly Val Ser Gly Gln Gly Lys Thr Asn Leu
    450   455   460

Leu Gly Asn Pro Asp Leu Lys Pro Glu Glu Ser Val Ser Tyr Glu Ala
465   470   475   480

Gly Val Tyr Tyr Asp Asn Pro Ala Gly Leu Asn Ala Asn Val Thr Gly
            485   490   495

Phe Met Thr Asp Phe Ser Asn Lys Ile Val Ser Tyr Ser Ile Asn Asp
            500   505   510

Asn Thr Asn Ser Tyr Val Asn Ser Gly Lys Ala Arg Leu His Gly Val
    515   520   525

Glu Phe Ala Gly Thr Leu Pro Leu Trp Ser Glu Asp Val Thr Leu Ser
    530   535   540

Leu Asn Tyr Thr Trp Thr Arg Ser Glu Gln Arg Asp Gly Asp Asn Lys
545   550   555   560

Gly Ala Pro Leu Ser Tyr Thr Pro Glu His Met Val Asn Ala Lys Leu
            565   570   575

Asn Trp Gln Ile Thr Glu Glu Val Ala Ser Trp Leu Gly Ala Arg Tyr
            580   585   590

Arg Gly Lys Thr Pro Arg Phe Thr Gln Asn Tyr Ser Ser Leu Ser Ala
        595   600   605

Val Gln Lys Lys Val Tyr Asp Glu Lys Gly Glu Tyr Leu Lys Ala Trp
    610   615   620

Thr Val Val Asp Ala Gly Leu Ser Trp Lys Met Thr Asp Ala Leu Thr
625   630   635   640

Leu Asn Ala Ala Val Asn Asn Leu Leu Asn Lys Asp Tyr Ser Asp Val
            645   650   655

Ser Leu Tyr Ser Ala Gly Lys Ser Thr Leu Tyr Ala Gly Asp Tyr Phe

Gln Thr Gly Ser Ser Thr Thr Gly Tyr Val Ile Pro Glu Arg Asn Tyr
      675   680   685

Trp Met Ser Leu Asn Tyr Gln Phe
    690   695

<210> 6
<211> 652
<212> PRT
<213> Escherichia coli

<400> 6
Met Ser Arg Phe Asn Pro Ser Pro Val Ser Leu Ser Val Thr Leu Gly
  1   5   10   15

Leu Met Phe Ser Ala Ser Ala Phe Ala Gln Asp Ala Thr Lys Thr Asp
              20   25   30

Glu Thr Met Val Val Thr Ala Ala Gly Tyr Ala Gln Val Ile Gln Asn
          35   40   45

Ala Pro Ala Ser Ile Ser Val Ile Ser Arg Glu Asp Leu Glu Ser Arg
      50   55   60

Tyr Tyr Arg Asp Val Thr Asp Ala Leu Lys Ser Val Pro Gly Val Thr
 65   70   75   80

Val Thr Gly Gly Gly Asp Thr Thr Asp Ile Ser Ile Arg Gly Met Gly
              85   90   95

Ser Asn Tyr Thr Leu Ile Leu Val Asp Gly Lys Arg Gln Thr Ser Arg
              100   105   110

Gln Thr Arg Pro Asn Ser Asp Gly Pro Gly Ile Glu Gln Gly Trp Leu
          115   120   125

Pro Pro Leu Gln Ala Ile Glu Arg Ile Glu Val Ile Arg Gly Pro Met
    130   135   140

Ser Thr Leu Tyr Gly Ser Asp Ala Ile Gly Gly Val Ile Asn Ile Ile
145   150   155   160

Thr Arg Lys Asp Gln Gln Gln Trp Ser Gly Asn Val Gln Leu Ser Thr
              165   170   175

Val Val Gln Glu Asn Arg Ala Ser Gly Asp Glu Gln Ser Ala Asn Phe
              180   185   190

Phe Val Thr Gly Pro Leu Ser Asp Ala Leu Ser Leu Gln Val Tyr Gly
          195   200   205

Gln Thr Thr Gln Arg Asp Glu Asp Glu Ile Glu His Gly Tyr Gly Asp
      210   215   220

Lys Ser Leu Arg Ser Leu Thr Ser Lys Leu Asn Tyr Gln Leu Asn Pro

-10-

```
          225                    230                    235                    240

    Asp His Gln Leu Gln Leu Glu Ala Gly Val Ser Ala Gln Asp Arg Glu
                  245                    250                    255

    Asn Asn Val Gly Lys Ser Ala Gln Ser Ser Gly Cys Arg Gly Thr Cys
                  260                    265                    270

    Ser Asn Thr Asp Asn Gln Tyr Arg Arg Asn His Val Ala Val Ser His
                  275                    280                    285

    Gln Gly Asp Trp Gln Gly Val Gly Gln Ser Asp Thr Tyr Leu Gln Tyr
        290                    295                    300

    Glu Glu Asn Thr Asn Lys Ser Arg Glu Met Ser Ile Asp Asn Thr Val
    305                    310                    315                    320

    Phe Lys Ser Thr Leu Val Ala Pro Ile Gly Glu His Met Leu Ser Phe
                  325                    330                    335

    Gly Val Glu Gly Lys His Glu Ser Leu Glu Asp Lys Thr Ser Asn Lys
                  340                    345                    350

    Ile Ser Ser Arg Thr His Ile Ser Asn Thr Gln Trp Ala Gly Phe Ile
                  355                    360                    365

    Glu Asp Glu Trp Ala Leu Ala Glu Gln Phe Arg Leu Thr Phe Gly Gly
        370                    375                    380

    Arg Leu Asp His Asp Lys Asn Tyr Gly Ser His Phe Ser Pro Arg Val
    385                    390                    395                    400

    Tyr Gly Val Trp Asn Leu Asp Pro Leu Trp Thr Val Lys Gly Gly Val
                  405 ·                  410                    415

    Ser Thr Gly Phe Arg Ala Pro Gln Leu Arg Glu Val Thr Pro Asp Trp
                  420                    425                    430

    Gly Gln Val Ser Gly Gly Gly Asn Ile Tyr Gly Asn Pro Asp Leu Gln
                  435                    440                    445

    Pro Glu Thr Ser Ile Asn Lys Glu Leu Ser Leu Met Tyr Ser Thr Gly
        450                    455                    460

    Ser Gly Leu Ala Ala Ser Leu Thr Ala Phe His Asn Asp Phe Lys Asp
    465                    470                    475                    480

    Lys Ile Thr Arg Val Ala Cys Pro Ala Asn Ile Cys Thr Ala Gly Pro
                  485                    490                    495

    Asn Gln Trp Gly Ala Thr Pro Thr Tyr Arg Val Asn Ile Asp Glu Ala
                  500                    505                    510

    Glu Thr Tyr Gly Ala Glu Ala Thr Leu Ser Leu Pro Ile Thr Glu Ser
                  515                    520                    525

    Val Glu Leu Ser Ser Ser Tyr Thr Tyr Thr His Ser Glu Gln Lys Ser
                  530                    535                    540
```

Gly Asn Phe Ala Gly Arg Pro Leu Leu Gln Leu Pro Lys His Leu Phe
545   550   555   560

Asn Ala Asn Leu Ser Trp Gln Thr Thr Asp Arg Leu Asn Ser Trp Ala
                565   570   575

Asn Leu Asn Tyr Arg Gly Lys Glu Met Gln Pro Glu Gly Gly Ala Ser
            580   585   590

Asn Asp Asp Phe Ile Ala Pro Ser Tyr Thr Phe Ile Asp Thr Gly Val
            595   600   605

Thr Tyr Ala Leu Thr Asp Thr Ala Thr Ile Lys Ala Ala Val Tyr Asn
    610   615   620

Leu Phe Asp Gln Glu Val Asn Tyr Ala Glu Tyr Gly Tyr Val Glu Asp
625   630   635   640

Gly Arg Arg Tyr Trp Leu Gly Leu Asp Ile Ala Phe
            645   650

<210> 7
<211> 663
<212> PRT
<213> Escherichia coli

<400> 7
Met Phe Arg Leu Asn Pro Phe Val Arg Val Gly Leu Cys Leu Ser Ala
1   5   10   15

Ile Ser Cys Ala Trp Pro Val Leu Ala Val Asp Asp Asp Gly Glu Thr
            20   25   30

Met Val Val Thr Ala Ser Ser Val Glu Gln Asn Leu Lys Asp Ala Pro
        35   40   45

Ala Ser Ile Ser Val Ile Thr Gln Glu Asp Leu Gln Arg Lys Pro Val
    50   55   60

Gln Asn Leu Lys Asp Val Leu Lys Glu Val Pro Gly Val Gln Leu Thr
65   70   75   80

Asn Glu Gly Asp Asn Arg Lys Gly Val Ser Ile Arg Gly Leu Asp Ser
            85   90   95

Ser Tyr Thr Leu Ile Leu Val Asp Gly Lys Arg Val Asn Ser Arg Asn
            100   105   110

Ala Val Phe Arg His Asn Asp Phe Asp Leu Asn Trp Ile Pro Val Asp
        115   120   125

Ser Ile Glu Arg Ile Glu Val Val Arg Gly Pro Met Ser Ser Leu Tyr
    130   135   140

Gly Ser Asp Ala Leu Gly Gly Val Val Asn Ile Ile Thr Lys Lys Ile
145   150   155   160

Gly Gln Lys Trp Ser Gly Thr Val Thr Val Asp Thr Thr Ile Gln Glu

-12-

```
                          165                    170                    175
His Arg Asp Arg Gly Asp Thr Tyr Asn Gly Gln Phe Phe Thr Ser Gly
            180                 185             190

Pro Leu Ile Asp Gly Val Leu Gly Met Lys Ala Tyr Gly Ser Leu Ala
            195                 200             205

Lys Arg Glu Lys Asp Asp Pro Gln Asn Ser Thr Thr Thr Asp Thr Gly
    210             215                 220

Glu Thr Pro Arg Ile Glu Gly Phe Ser Ser Arg Asp Gly Asn Val Glu
225             230             235                     240

Phe Ala Trp Thr Pro Asn Gln Asn His Asp Phe Thr Ala Gly Tyr Gly
                245             250                     255

Phe Asp Arg Gln Asp Arg Asp Ser Asp Ser Leu Asp Lys Asn Arg Leu
            260             265             270

Glu Arg Gln Asn Tyr Ser Val Ser His Asn Gly Arg Trp Asp Tyr Gly
            275             280             285

Thr Ser Glu Leu Lys Tyr Tyr Gly Glu Lys Val Glu Asn Lys Asn Pro
    290             295             300

Gly Asn Ser Ser Pro Ile Thr Ser Glu Ser Asn Thr Val Asp Gly Lys
305             310             315                     320

Tyr Thr Leu Pro Leu Thr Ala Ile Asn Gln Phe Leu Thr Val Gly Gly
            325             330             335

Glu Trp Arg His Asp Lys Leu Ser Asp Ala Val Asn Leu Thr Gly Gly
            340             345             350

Thr Ser Ser Lys Thr Ser Ala Ser Gln Tyr Ala Leu Phe Val Glu Asp
    355             360             365

Glu Trp Arg Ile Phe Glu Pro Leu Ala Leu Thr Thr Gly Val Arg Met
    370             375             380

Asp Asp His Glu Thr Tyr Gly Glu His Trp Ser Pro Arg Ala Tyr Leu
385             390             395                     400

Val Tyr Asn Ala Thr Asp Thr Val Thr Val Lys Gly Gly Trp Ala Thr
            405             410             415

Ala Phe Lys Ala Pro Ser Leu Leu Gln Leu Ser Pro Asp Trp Thr Ser
            420             425             430

Asn Ser Cys Arg Gly Ala Cys Lys Ile Val Gly Ser Pro Asp Leu Lys
            435             440             445

Pro Glu Thr Ser Glu Ser Trp Glu Leu Gly Leu Tyr Tyr Met Gly Glu
    450             455             460

Glu Gly Trp Leu Glu Gly Val Glu Ser Ser Val Thr Val Phe Arg Asn
465             470             475                     480
```

-13-

```
Asp Val Lys Asp Arg Ile Ser Ile Ser Arg Thr Ser Asp Val Asn Ala
                485                 490                 495

Ala Pro Gly Tyr Gln Asn Phe Val Gly Phe Glu Thr Gly Ala Asn Gly
                500                 505                 510

Arg Arg Ile Pro Val Phe Ser Tyr Tyr Asn Val Asn Lys Ala Arg Asn
            515                 520                 525

Gln Gly Val Glu Thr Glu Leu Lys Ile Pro Phe Asn Asp Glu Trp Lys
        530                 535                 540

Leu Ser Ile Asn Tyr Thr Tyr Asn Asp Gly Arg Asp Val Ser Asn Gly
545                 550                 555                 560

Glu Asn Lys Pro Leu Ser Asp Leu Pro Phe His Leu Ala Leu Glu Asp
                565                 570                 575

Trp Ser Phe Tyr Val Ser Gly His Tyr Thr Gly Gln Lys Arg Ala Asp
                580                 585                 590

Ser Ala Thr Ala Lys Thr Pro Gly Gly Tyr Thr Ile Trp Asn Thr Gly
            595                 600                 605

Ala Ala Trp Gln Val Thr Lys Asp Val Lys Leu Arg Ala Gly Val Leu .
    610                 615                 620

Asn Leu Gly Asp Lys Thr Ala Asn Gly Thr Leu Asp Trp Lys Pro Asp
625                 630                 635                 640

Leu Ser Arg Asp Asp Tyr Ser Tyr Asn Glu Asp Gly Arg Arg Tyr Phe
                645                 650                 655

Met Ala Val Asp Tyr Arg Phe
                660


<210> 8
<211> 32
<212> DNA
<213> Escherichia coli

<400> 8
ggggatccaa ttctggcatg ccgaggcagt cg                              32


<210> 9
<211> 33
<212> DNA
<213> Escherichia coli

<400> 9
ggaccgcctt gtcaccgttg ctcttagatc tgg                             33


<210> 10
<211> 34
<212> DNA
<213> Escherichia coli
```

<400> 10
ggaaggatcc ccgaacacgc catacggata gctg   34

<210> 11
<211> 35
<212> DNA
<213> Escherichia coli

<400> 11
gcaacggtga cgttgaggac cgccagatct aaagg   35

<210> 12
<211> 300
<212> DNA
<213> Escherichia coli

<400> 12
atggaagcag caaatttaag tccttctggt gcagtaatgc cgctggcgac ctcactcagt 60
ggaaataact cagtggatga gaagacagga gtgattaaac cagaaaatgg aacaaatcgc 120
accgttagag ttatagccgg attagcactt accactacgg ctctggcagc tctaggtaca 180
ggtattgcag cggcatgctc ggagacgagc agcacagaat acttagccct gggtattact 240
tctggcgtac taggtactct tactgcggtt ggcggtgcat tagcgatgaa atatgcctaa 300